Chapter 17: Negative Binomial RVs

Meike Niederhausen and Nicky Wakim

2023-10-30

Table of contents

Learning Objectives

1. Identify the variable and the parameters of a Negative Binomial distribution in a word problem, and state what the variable and parameters mean.
2. Use the formulas for the pmf/CDF, expected value, and variance to answer questions and find probabilities.

Properties of Negative Binomial RVs

★ good NB example in regression analysis

- **Scenario:** There are repeated independent trials, each resulting in a success or failure, with constant probability of success for each trial. We are counting the number of trials until the r^{th} success.
- Shorthand: $X \sim \text{NegBin}(p, r)$ or $X \sim \text{NB}(p, r)$
- Negative binomial is sum of r geometric distributions

$$X = \text{Number of independent trials until } r^{th} \text{ success}$$

$$p_X(x) = P(X = x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \text{ for } x = r, r+1, r+2, \ldots$$

$$E(X) = \frac{r}{p}$$

$$Var(X) = \frac{rq}{p^2} = \frac{r(1-p)}{p^2}$$

x-r failures
r successes

possible combos of 1st x-1 throws w/ the r-1 successes

1 → trial
bern
Σ ↓
Binom
n trials

1 → # trials until 1 success
geom
↓ Σ
NB
r successes (over as many trials needed)

Hitting more than 1 bullseye

Example 1

Consider again the bullseye example, where we throw darts at a dartboard until we hit the bullseye. Assume throws are independent and the probability of hitting the bullseye is 0.01 for each throw.

1. What is the expected value and variance of the number of throws needed to hit 5 bullseyes?

let X = # independent throws until (including) 5 successes

$r = 5$

$$E(X) = \frac{r}{p} = \frac{5}{0.01} = \underline{500}$$

Y_i = # trials our 1st (or next) bullseye

$Y_i \overset{iid}{\sim} Geo(p = 0.01)$ for $i = 1, 2, 3, 4, 5$ ← r

Y_i is ind & identically distributed as...

$$E(X) = E\left[\sum_{i=1}^{5} Y_i\right] = \sum_{i=1}^{5} E(Y_i) = 5E(Y_1) = 5\left(\frac{1}{0.01}\right)$$

(linearity) (iid)

$$Var(X) = Var\left(\sum_{i=1}^{r} Y_i\right) = \sum_{i=1}^{r} Var(Y_i) = r\,Var(Y_1) = r\left[\frac{1-p}{p^2}\right] = \frac{5(0.99)}{0.01^2} = 49500$$

X: NB; $\sum Y_i$: sum geo

Hitting more than 1 bullseye

Example 1

Consider again the bullseye example, where we throw darts at a dartboard until we hit the bullseye. Assume throws are independent and the probability of hitting the bullseye is 0.01 for each throw.

2. What is the probability that the 5^{th} bullseye is on the 20^{th} throw?

$5^{\text{th}} \rightarrow r = 5$

$20^{\text{th}} \text{ throw} \rightarrow P(X = 20)?$

$$p_X(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \quad \text{for } x = r, r+1, r+2, \ldots$$

$$p_X(20) = P(X=20) = \binom{20-1}{5-1}(0.99)^{20-5}(0.01)^5$$

$\binom{20-1}{5-1}$: # ways to arrange 1st 19 throws w/ 4 successes (20th throw is definitely a success)

$= 0.00000033$

Y = # throws prior (excluding rth success)

$\hookrightarrow Y = X - 1$